Chapter 9

General discussion

Proteomics

The successful completion of the human genome project had led to a tremendous increase in our understanding of the molecular basis of diseases. However, a comprehensive understanding of the dynamic protein pathways involved in normal and disease states, and in response to medical treatment, is required if we want to effectively treat diseases. The next major challenge towards this aim is to identify the constituents of the human proteome in order to understand the human genome[1]. Advances in proteomics technology offer great promise in the understanding and treatment of the molecular basis of disease. Proteome analysis is based on the combination of three technologies: a resolutive method for separating proteins as a function of different physicochemical criteria (e.g., mass, isoelectric point and hydrophobicity), mass spectrometry and bioinformatics tools[2].
The true scientific goal of proteomic pattern analysis is in fact biomarker discovery. However, since the study by Petricoin et al.[3] on proteomic patterns to detect ovarian cancer, the use of Surface-Enhanced Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (SELDI-TOF-MS) protein profiling as a diagnostic tool, has become an important subject of investigation too[4]. Until now, this approach has been suggested for different diseases, like ovarian[3, 5-9], prostate[10-14] and lung[15] cancer, but also for inflammatory diseases [16, 17]. This thesis is especially focused on the SELDI-TOF-MS technique.
Recently, Poon[18] described some major limitations of the SELDI-TOF-MS technology. One important issue is that the limited sensitivity causes difficulty in the identification of potential diagnostic proteins present at concentrations below the mg/l level in serum and/or plasma. To date, this technology has not yet identified protein markers present at the µg/l level. Unfortunately, the dynamic range of serum/plasma proteins covers ten orders of magnitude (low ng/l to g/l) in abundance. To overcome limited detection sensitivity, fractionation by anion-exchange chromatography by stepwise pH gradient elution can be performed. Other fractionation technologies, like isoelectric focusing and reverse phase chromatography can also be applied to the fractionation step. Alternately, removal of the major abundant serum proteins can also help to identify low-abundance proteins.
Another important issue is that most of the diagnostic SELDI-TOF-MS peaks were found to be acute-phase proteins, their fragments or isoforms. Diamandis[19] and others[20-22] indicated that a very large number of candidate cancer biomarkers previously identified using the SELDI-TOF-MS technology are liver-derived products originating from common disease mechanisms, like infection or cachexia resulting in for instance an acute phase response. Unfortunately until now most studies present biomarkers related to these non-specific common disease mechanisms.
Diamandis also described that it is strange that in many cancer studies a large number of discriminatory peaks are found which are down-regulated in the cancer

population, in comparison to normal controls. He indicated that there are currently no useful markers whose concentration is actually decreased in the serum of patients with cancer. The most useful cancer biomarkers originate from tumor cells and their concentration is increased in the circulation and correlates with tumor burden. A possible gene down-regulation effect in tumor cells would have been highly speculative unless the marker under discussion has absolute tissue specificity and the tumor overtakes the normal tissue, thus reducing the marker concentration in serum. A more likely explanation is that these molecules represent highly abundant proteins produced by the liver or other organs, whose concentration is decreased due to cancer cachexia or malnutrition, non-specific effects for many cancer types, as described before. Another important aspect put forward by Diamandis is that it is known that there is a proportional relationship between biomarker concentration and tumor stage. In most of the published data, such a relation was not found.

There is no doubt that SELDI-TOF-MS is a useful tool in biomedical research, particularly in biomarker discovery. Since reliability and stability of the ProteinChip arrays have been questioned worldwide, improvement of their quality and stability, or providing the proof of negligible batch-to-batch variations in their quality are needed urgently. Systematic investigations of the effects of various experimental variables on the SELDI-TOF-MS proteomic profiles are needed. With reliable ProteinChip arrays, optimized standard operating procedures, appropriate study design and experimental precautions, the proteomic profiling results should become more reliable[18].

The importance of the pre-and post-analytical strategies were reported in chapter 2 of this thesis. This overview clearly underlines the need for better standardization and careful description of the methods including technical details in all future studies. It is essential that sample collection from both the patient and control populations is completely identical and accurately standardized. A standard protocol for the collection of serum samples according to World Healthy Organization (WHO) recommendations on “Anticoagulants in diagnostic laboratory investigations” (2002) is suggested in this thesis. There are other international organizations, such as the International Federation of Clinical Chemistry (IFCC) and the Human Proteome Organization (HUPO) that are looking into some standardization issues.

In chapter 3 was indicated that any new technology, particularly one being presented as a potential diagnostic tool, requires stringent quality control to evaluate analytical performance over time. Instrument performance must be compared not only during one experiment, but also over the course of time. Therefore a standard protocol for calibration of the Matrix Assisted Laser Desorption/Ionization-Time of Flight-Mass Spectrometry (MALDI-TOF-MS) part of the Protein Biosystem IIc (PBS IIc) instrument was defined and acceptance criteria for the independent certified quality control (QC) samples were established in this

thesis. The acceptance criteria can be exceeded because of different factors, like errors during preparations and handling of the calibration or QC sample as well as instrumental problems. Because an error introduced during processing can be difficult or impossible to trace once the experiment is completed. It is best to rigorously control the experimental procedure to minimize the introduction of variation in the first place

Besides biomarker discovery, SELDI-TOF-MS can also be used in basic research, for example for discovery of protein-protein interactions[23]. ProteinChip arrays can be coated with a particular DNA sequence to examine the interaction between a DNA molecule and transcription factors[24]. With coating of specific antibodies on the binding surface of the ProteinChip arrays, the SELDI-TOF-MS technology can be used as a specific novel immunoassay for the quantification of different variants of the target protein[25].

Detection of potential biomarkers

This thesis contains a review (Chapter 4) about proteomics studies in multiple sclerosis, rheumatic diseases and lung inflammatory diseases. A lot of proteins were identified and in some of the studies the identified proteins were also validated with other tests. The heat shock proteins were entitled as biomarkers with potential for further research in multiple sclerosis. Myeloid-Related protein 8 MRP-8 was found in three different rheumatoid arthritis studies with different sample materials and could be a potential marker for rheumatoid arthritis. Alpha1-antitrypsin was validated in two studies as a marker for sarcoidosis, although in one study sarcoidosis patients were compared with idiopathic pulmonary fibrosis patients and in the other study the sarcoidosis patients were compared with healthy control samples. Alpha1-antitrypsin was also found as a marker for cystic fibrosis (CF), together with myeloperoxidase and immunoglobulin G (IgG). This thesis also contains own data regarding biomarker finding with proteomics in two chronic inflammatory diseases and in lacunar stroke. The aim of the study described in chapter 5 was the detection of potential biomarkers for sarcoidosis using SELDI-TOF-MS. The markers obtained with SELDI-TOF-MS resulted in higher sensitivity and specificity compared to the routinely used markers angiotensin converting enzyme (ACE) and soluble Interleukin-2 Receptor (sIL-2R). We were able to identify one peak as the alpha-2 chain of haptoglobin. The clinical problems associated with sarcoidosis might be related to a decreased anti-oxidant defense [26]. Glucose-6-phosphate dehydrogenase (G6PD) is involved in the anti-oxidant defence. It is known that in the presence of oxidative stress, G6PD deficiency leads to hemolysis[27], which causes a decrease in haptoglobin, which as such might clarify the lower haptoglobin fragment peaks in sarcoidosis patients.

The aim of the study presented in chapter 6 was to detect potential biomarkers for ankylosing spondylitis (AS). Sera of AS patients and healthy controls were used to detect potential biomarkers for the diagnosis of AS. Two multimarkers were found which could classify the AS and healthy control samples. Not all peaks could be identified, but two peaks were both successfully identified as apolipoprotein A-I (ApoA1).
In Chapter 7 a study was described to detect potential biomarkers for lacunar stroke. SELDI-TOF-MS, followed by protein identification, was performed in samples of first-ever lacunar stroke patients with magnetic resonance imaging showing a single symptomatic lacunar lesion (type I), and samples of patients with multiple additional “silent” lesions and extensive white matter lesions (type II). A 16 kDa protein, identified as alpha-2-chain of haptoglobin, was up-regulated in type I compared to type II. The up-regulation of the alpha-2-chain in type I is related to a higher haptoglobin-2 allele frequency. The phenotype distribution deviates towards a high haptoglobin-1 allele frequency in both groups compared to population reference, suggesting a role for the haptoglobin gene in the etiology of cerebral small vessel disease. The even higher haptoglobin-1 allele frequency in type II than in type I implies a promoting role for haptoglobin-1 in developing multiple silent lacunar lesions and as a result white matter lesions. The haptoglobin-1 allele frequency could be a reflection of a difference in underlying vascular pathology between the two types and the influence of the haptoglobin phenotype may thus be relevant in predisposing to one type of small vessel pathology more than the other.
Both the results of the review article as our own proteomics studies illustrate that especially biomarkers associated with acute-phase reaction and biomarkers with a relatively high serum concentration were detected and identified. The biomarkers, which are presented in this thesis, need to be validated in the future by using blind sample sets. To verify the robustness of the potential markers, samples from different institutions should also be included, to check if the sensitivities and specificities could be reproduced.

Glycomics

The analytical techniques described above for finding biomarkers can also be used for more extensive analysis of proteins, like protein glycosylation. In this thesis, we looked for glycosylation defects in congenital disorders of glycosylation (CDG) and galactosemia. The hypothesis is that glycan abnormalities of IGFBP-3 can be the cause of the diminished bone mass and growth in classic galactosemia. The aim of this study was to detect possible differences in IGFBP-3 isoforms (intact protein and fragments) in patients with galactosemia, CDG and healthy controls. Chapter 8 provides detailed information about the optimization of the immunoprecipitation method used in this study to purify and isolate IGFBP-3 from serum. Because the

glycosylation defect in CDG type Ia is known, we started to compare CDG type Ia samples with healthy controls. CDG-Ia is due to phosphomannomutase deficiency[28], a key enzyme in the synthesis of guanosine 5'-diphosphate-D-mannose which is required for N-glycan assembly. Deficient synthesis of N-glycans results in a deficient incorporation of sialic acid, the terminal negatively charged sugar, so that glycoprotein molecules acquire a more positive charge, which causes a shift in the IEF patterns from more acidic to more basic[29]. Abnormal isoelectric focusing of transferrin in CDG with a shift towards the cathode was described earlier by Jaeken et al. They found these features in CDG[30] and in untreated classic galactosemia[31]. IGFBP-3 has three N-glycosylation sites and disturbed glycosylation results in a cathodal shift[32]. By comparing the serum samples of paediatric CDG type Ia patients and control serum of paediatric healthy persons, we could detect a shift from more acidic to more basic isoforms for the IGFBP-3 protein in CDG type Ia. This is in agreement with the transferrin results. By using CDG type Ia samples we were able to clinically validate our method and in future studies we will enlarge our sample group and we will also investigate the isoform patterns of IGFBP-3 in galactosemia to confirm that glycan abnormalities of IGFBP-3 can be the cause of the diminished bone mass and growth in classic galactosemia.

Directions for future research

Because the limited sensitivity of the SELDI-TOF-MS causes difficulties in the identification of potential diagnostic proteins present at concentrations below the mg/l level, advancements in the MS instrumentation part of the SELDI-TOF-MS technology or similar technologies, like tandem mass spectrometry (MS/MS), are needed and will lead to the identification of potential biomarkers at µg/l or even lower level. For example, combination of SELDI-TOF-MS technology and MALDI tandem TOF-MS can offer much greater resolution and sensitivity and it will give the potential for direct protein identification. Combined use of *m/z* values and N-/C-terminal sequence tags can provide unambiguous identities of the SELDI-TOF-MS peaks[18].

Currently, the pipeline from translation of new biomarkers into diagnostic tests appears to have a bottleneck. A number of technical obstacles, like the limited sensitivity and relative low reproducibility, remain before routine proteomic analysis can be achieved in the clinic. Standardization of methodologies and dissemination of proteomic data into publicly available databases is starting to overcome part of these hurdles. At present the most promising application for proteomics is the detection of specific subsets of protein biomarkers for certain diseases, rather than large scale full protein profiling[1]. As described before, this thesis is especially focused on the SELDI-TOF-MS technology. For future proteomic profiling studies with the SELDI-TOF-MS technology it is important to examine the effect of sample collection and storage conditions on the data, but also other variables that could potentially bias the data, either biologically of bioinformatically. It is also essential to identify the discriminative peaks successfully and find a link with the disease pathobiology. In cancer studies, the peak intensities should correlate with the tumor burden or tumor stage. The SELDI-TOF-MS technology can also be used for purifying proteins. Antibodies can be coupled on the ProteinChip arrays.

For the IGFBP-3 study, negative images of two-dimensional gel electrophoresis (2-DE) Western blots of paediatric galactosemic patients, CDG patients and healthy controls will be compared and the (glyco)proteins in the discriminating spots will be identified with MALDI-TOF-MS.

References

1. Verrills NM. Clinical proteomics: present and future prospects. Clin Biochem Rev 2006;27:99-116.
2. Solassol J, Jacot W, Lhermitte L, Boulle N, Maudelonde T, Mange A. Clinical proteomics and mass spectrometry profiling for cancer detection. Expert Rev Proteomics 2006;3:311-320.
3. Petricoin III EF, Mills GB, Kohn EC, Liotta LA. Proteomic patterns in serum and identification of ovarian cancer. Lancet 2002;360:170-171.
4. Petricoin EF, Ardekani AM, Hitt BA, Levine PJ, Fusaro VA, Steinberg SM, Mills GB, Simone C, Fishman DA, Kohn EC, Liotta LA. Use of proteomic patterns in serum to identify ovarian cancer. Lancet 2002;359:572-577.
5. Kozak KR, Amneus MW, Pusey SM, Su F, Luong MN, Luong SA, Reddy ST, Farias-Eisner R. Identification of biomarkers for ovarian cancer using strong anion-exchange ProteinChips: potential use in diagnosis and prognosis. Proc Natl Acad Sci U S A 2003;100: 12343-12348.
6. Zhang Z, Bast RC Jr, Yu Y, LiJ, Sokoll LJ, Rai AJ, Rosenzweig JM, Cameron B, Wang YY, Meng XY, Berchuck A, Van Haaften-Day C, Hacker NF, de Bruijn HW, van der Zee AG, Jacobs IJ, Fung ET, Chan DW. Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer. Cancer Res 2004;64:5882-5890.
7. Vlahou A, Schorge JO, Gregory BW, Coleman RL. Diagnosis of Ovarian Cancer Using Decision Tree Classification of Mass Spectral Data. J Biomed Biotechnol 2003;2003:308-314.
8. Ye B, Cramer DW, Skates SJ, Gygi SP, Pratomo V, Fu L, Horick NK, Licklider LJ, Schorge JO, Berkowitz RS, Mok SC. Haptoglobin-alpha subunit as potential serum biomarker in ovarian cancer: identification and characterization using proteomic profiling and mass spectrometry. Clin Cancer Res 2003;9:2904-2911.
9. Rai AJ, Zhang Z, Rosenzweig J, Shih Ie M, Pham T, Fung ET, Sokoll LJ, Chan DW. Proteomic approaches to tumor marker discovery. Arch Pathol Lab Med 2002;126:1518-1526.
10. Banez LL, Prasanna P, Sun L, Ali A, Zou Z, Adam BL, McLeod DG, Moul JW, Srivastava S. Diagnostic potential of serum proteomic patterns in prostate cancer. J Urol 2003; 170:442-446.
11. Adam BL, Qu Y, Davis JW, Ward MD, Clements MA, Cazares LH, Semmes OJ, Schellhammer PF, Yasui Y, Feng Z, Wright GL Jr. Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men. Cancer Res 2002;62:3609-3614.
12. Petricoin EF 3rd, Ornstein DK, Paweletz CP, Ardekani A, Hackett PS, Hitt BA, Velassco A, Trucco C, Wiegand L, Wood K, Simone CB, Levine PJ, Linehan WM, Emmert-Buck MR, Steinberg SM, Kohn EC, Liotta LA. Serum proteomic patterns for detection of prostate cancer. J Natl Cancer Inst 2002;94:1576-1578.
13. Qu Y, Adam BL, Yasui Y, Ward MD, Cazares LH, Schellhammer PF, Feng Z, Semmes OJ, Wright GL Jr. Boosted decision tree analysis of surface-enhanced laser desorption/ionization mass spectral serum profiles discriminates prostate cancer from noncancer patients. Clin Chem 2002;48: 1835-1843.
14. Li J, White N, Zhang Z, Rosenzweig J, Mangold LA, Partin AW, Chan DW. Detection of prostate cancer using serum proteomics pattern in a histologically confirmed population. J Urol 2004;171:1782-1787.
15. Zhukov TA, Johanson RA, Cantor AB, Clark RA, Tockman MS. Discovery of distinct protein profiles specific for lung tumors and pre-malignant lung lesions by SELDI mass spectrometry. Lung Cancer 2003;40:267-279.
16. Poon TC, Hui AY, Chan HL, Ang IL, Chow SM, Wong N, Sung JJ. Prediction of Liver Fibrosis and Cirrhosis in Chronic Hepatitis B Infection by Serum Proteomic Fingerprinting: A Pilot Study. Clin Chem 2004;51:328-335.
17. Zhu XD, Zhang WH, Li CL, Xu Y, Liang WJ, Tien P. New serum biomarkers for detection of HBV-induced liver cirrhosis using SELDI protein chip technology. World J Gastroenterol 2004; 10:2327-2329.
18. Poon TC. Opportunities and limitations of SELDI-TOF-MS in biomedical research: practical advices. Expert Rev Proteomics 2007;4:51-65.

19. Diamandis EP. Serum Proteomic Profiling by Matrix-Assisted Laser Desorption-Ionization Time-of-Flight Mass Spectrometry for Cancer Diagnosis: Next Steps. Cancer Res 2006;66:5540-5541.
20. Le L, Chi K, Tyldesley S, Flibotte S, Diamond DL, Kuzyk MA, Sadar MD. Identification of serum amyloid A as a biomarker to distinguish prostate cancer patients with bone lesions. Clin Chem 2005;51:695-707.
21. Lee IN, Chen CH, Sheu JC, Lee HS, Huang GT, Chen DS, Yu CY, Wen CL, Lu FJ, Chow LP. Identification of complement C3a as a candidate biomarker in human chronic hepatitis C and HCV-related hepatocellular carcinoma using a proteomics approach. Proteomics 2006;6: 2865-2873.
22. Engwegen JY, Helgason HH, Cats A, Harris N, Bonfrer JM, Schellens JH, Beijnen JH. Identification of serum proteins discriminating colorectal cancer patients and healthy controls using surface-enhanced laser desorption ionisation-time of flight mass spectrometry. World J Gastroenterol 2006;12:1536-1544.
23. Howell JM, Winstone TL, Coorssen JR, Turner RJ. An evaluation of in vitro protein-protein interaction techniques: assessing contaminating background proteins. Proteomics 2006;6: 2050-2069.
24. Bane TK, LeBlanc JF, Lee TD, Riggs AD. DNA affinity capture and protein profiling by SELDI-TOF mass spectrometry: effect of DNA methylation. Nucleic Acids Res 2002;30:e69.
25. Schweigert FJ, Wirth K, Raila J. Characterization of the microheterogeneity of transthyretin in plasma and urine using SELDI-TOF-MS immunoassay. Proteome Sci 2004;2:5.
26. Rothkrantz-Kos S, Drent M, Vuil H, De Boer M, Bast A, Wouters EF, Roos D, van Dieijen-Visser MP. Decreased redox state in red blood cells from patients with sarcoidosis. Sarcoidosis Vasc Diffuse Lung Dis 2002;19:114-120.
27. Dhaliwal G, Cornett PA, Tierney LM Jr. Hemolytic anemia. Am Fam Physician 2004;69: 2599-2606.
28. Van Schaftingen E, Jaeken J. Phosphomannomutase deficiency is a cause of carbohydrate-deficient glycoprotein syndrome type I. FEBS Lett 1995;377:318-320.
29. Grunewald S, Matthijs G, Jaeken J. Congenital disorders of glycosylation: a review. Pediatr Res 2002;52:618-624.
30. Jaeken J, van Eijk HG, van der Heul C, Corbeel L, Eeckels R, Eggermont E. Sialic acid-deficient serum and cerebrospinal fluid transferrin in a newly recognized genetic syndrome. Clin Chim Acta 1984;144:245-247.
31. Jaeken J, Kint J, Spaapen L. Serum lysosomal enzyme abnormalities in galactosaemia. Lancet 1992;340:1472-1473.
32. Firth SM, Baxter RC. Characterisation of recombinant glycosylation variants of insulin-like growth factor binding protein-3. J Endocrinol 1999;160:379-387.